Brave New FRBR World

(Version 3)

Prepared for the 3rd IFLA Meeting of Experts on an International Cataloguing Code (IME ICC 3), December 12-14, 2005, Cairo, Egypt

Patrick Le Bœuf, Bibliothèque nationale de France
revised by Barbara Tillett, Library of Congress
delivered in Cairo by Elena Escolano Rodríguez, Biblioteca nacional de España

Introduction

You have just heard about the Paris Principles and ISBDs; I now would like to tell you about the FRBR model: what it is and what it is *not;* what it does and what it does *not;* and how it relates to our major topics during the present Meeting.

What FRBR is

— a model developed for IFLA

FRBR is the result of a study about the *functional requirements for bibliographic records* undertaken from 1992 through 1997 by a group of experts and consultants and commissioned by IFLA. The aim of that study "was to produce a *framework* that would provide a clear, precisely stated, and *commonly shared understanding* of what it is that the bibliographic record *aims to provide information about,* and what it is that we expect the record to achieve in terms of *answering user needs."*

— a reference model

FRBR is a reference model. The very words quoted above from the *FRBR Final Report* and emphasised by me make it very clear: it is but a *framework for commonly shared understanding.* It allows us to have the same structure in mind and to refer to the same concepts under the same appellations. It allows us to compare data that may happen not to be structured the same way.

— an E-R model

FRBR is an entity-relationship model. It defines a number of general classes ("entities") of things that are deemed relevant in the specific context of a library catalogue, a row of characteristics ("attributes") that pertain to each of these general classes, and the relationships that can exist between instances of these various classes.

The very core of FRBR consists in a group of 4 entities that pertain to documents themselves (the "things" that are being catalogued), from carrier to content. These 4 entities highlight the 4 distinct meanings attached to the word "kitab" (book). Originally, "kitab" just meant "something written," but there are at least four distinct ways to understand what the phrase "something written" covers:

– a "kitab" can be a physical object on which a text is written ("I have lost my *Kitab al-Fihrist"*); this first interpretation has to do with the notion of "copy," "exemplar;" the FRBR model uses the term "Item" to refer to that meaning.

– a "kitab" can be a set of physical objects on which the same text is written ("I have to order *Kitab al-Fihrist"*); this second interpretation is close to the notion of "publication;" but the FRBR model uses the more generic term "Manifestation," in order to cover unpublished materials as well.

– a "kitab" can be the text that is written ("Is this text *Kitab al-Fihrist* version one or *Kitab al-Fihrist* version two?"*); This third interpretation is equivalent to the notion of "text;" but once again the FRBR model uses a more generic term, "Expression," in order to refer to non-linguistic contents, such as music, maps, images, etc.; or

– a "kitab" can be the ideas that are expressed in the text that is written ("Ibn al-Nadim is the author of both versions of *Kitab al-Fihrist").* This fourth interpretation establishes the relationship between a text and its translations, and between the different versions of some human creation that is regarded as being "the same," in spite of the existence of its various versions; the FRBR model refers to such a notion through the term "Work."

So in the FRBR model we have a first group of entities: work, expression, manifestation, item.

A second group of entities comprises the 2 categories of entities that can create a Work, realise an Expression, produce or order a Manifestation, modify or possess an Item: Person and Corporate Body.

A third group of entities, reflecting what a Work may be *about,* comprises all of the above, plus 4 other entities that can *only* serve to express the subject of a Work: Concept, Object, Event, and Place.

Each of all of those entities is characterised by a number of "attributes" – i.e., the information elements that serve to identify the entity. For instance, a published Manifestation is characterised, among other information elements, by its date of publication; a Person is characterised by the name associated with that Person.

The FRBR model defines also relationships between those entities – as is natural enough for an entity-relationship model:

The relationships between a Work and one of its Expressions, between an Expression and one of its Manifestations, and between a Manifestation and one of its Items are said to be "structural" because the overall coherence of a bibliographic record depends on such relationships. They are reflected in our catalogues by the physical unity of descriptive records, by the link between the bibliographic record and holdings information, and sometimes by uniform title headings.

There are also "responsibility relationships" between Group 2 entities (Person and Corporate Body) and any of the entities belonging to Group 1. Such relationships are reflected in our catalogues by author headings.

And there are "subject relationships" between any entity in the model and the sole entity "Work" – that is to say: a Work is about another Work, or about an Item, or about a Corporate Body, or about a Place, etc. Such relationships are reflected in our catalogues by subject

headings. Those first three kinds of relationships are the most "basic" ones, deemed indispensable in library catalogues.

There are still other, more "subtle" kinds of relationships, such as whole/part relationships between two distinct Works or two distinct Expressions of the same Work, or intellectual relationships between two distinct Works, or between two distinct Expressions of the same Work or of distinct Works. Whole/part relationships exist also between two distinct Manifestations or two distinct Items of the same Manifestation; and the model defines "reproduction relationships" between two distinct Manifestations, or between two distinct Items of the same Manifestation, or between a Manifestation and an Item of a distinct Manifestation. All such relationships can be reflected in our catalogues as actual active links among records (bibliographic, authority, and/or holdings records), enabling end-users to navigate seamlessly the catalogue, or relationships can be indicated through textual notes that provide the information (like contents notes or history notes) but sometimes constrain end-users to launch a new query in order to navigate the catalogue and attain the "target" of the described relationship.

What FRBR is *not*

— a data model

Can FRBR be labelled a "data model?" It seems that the attributes it defines for each entity are in many cases too generic to allow for an implementation of the model such as it stands, without having to refine it. *Titles* for instance may have different natures; FRBR defines a Title attribute for each of the 3 entities Work, Expression, and Manifestation, but this categorisation of the "title notion" does not suffice to cover the typology of titles we actually need and currently use.

— an ISBD

Can FRBR be labelled a "new kind of ISBD?" No, roughly for the same reasons: FRBR does not state *how* to structure data elements nor *how* to display them so that their structure can be understood by just reading the description. Instead, FRBR provides an intellectual framework to typify data elements and to show how they are interrelated among distinct records.

— a cataloguing code

Similarly, FRBR cannot be said to be a cataloguing code. FRBR is not prescriptive, and does not tell you how you should record bibliographic information in day-to-day practice. FRBR is located at a merely conceptual level. Which does not mean, however, that FRBR has no *practical* utility. We now have a number of systems that actually work and are based on the way FRBR views bibliographic information. And there is a possibility that the future International Cataloguing Code be informed by the concepts highlighted by the model.

Some examples of implementation

What does the phrase "FRBR implementation" mean? I have just said that FRBR was *not* a data model, so how could it be "implemented?" At best, by designing an intermediate data model, based on it; at worst, by just *mistaking* it for a data model; in any case, by mapping either an extant format to FRBR, or FRBR to a new format.

AustLit Gateway

AustLit Gateway was the earliest database fully implementing FRBR. It is an atypical experiment, in that it applies to an exclusively literary corpus of Australian texts, and in that it results from the merging of a range of various, heterogeneous datasets, some of which were *not* based on ISBDs. It is not a catalogue, but rather a database aiming to provide scholars and students with as much information as possible about Australian writers and Australian literary works. As such, it is work-centred and it displays for each work all of its expressions and manifestations on a single Web page, instead of presenting users with rows of distinct bibliographic records, as we do in our current library catalogues.

Virtua

VTLS Inc. released in 2002 version 41.0 of the Virtua library system. For the very first time, a vendor made it possible for any library to create its own "FRBR catalogue." Extant MARC records can be "split" into the 4 levels of the FRBR Group 1 of entities, and any cataloguer can decide to account for bibliographic families rather than isolated documents, thanks to the FRBR structure. Virtua allows "flat records" and "FRBR records" to live side-by-side. The pattern followed when "splitting" records is based on Tom Delsey's mapping of MARC21 to FRBR. However, the cataloguing paradigm is still based on ISBDs — the "Manifestation level record" is not substantially different from any "traditional" ISBD-based record —, and the cataloguing format is still basically a MARC format, even though this MARC format is stored encapsulated in XML within the system, without cataloguers being aware of that.

On this fictitious example, borrowed from a presentation by Claude Détienne on the occasion of the 24th International Conference of MELCOM (Middle East Librarians Committee) in 2002, you can see, on the left part of the screen, the characteristic "FRBR tree," showing the uniform title for the Work *Alf layla wa-layla,* two of its linguistic Expressions, and brief records for three of its Manifestations, and, on the right part of the screen, a complete bibliographic record for one of those three Manifestations, with a link to information about one Item.

OCLC

OCLC has invested considerable effort in exploring the FRBR potential for very large bibliographic databases. OCLC has developed an algorithm that allows "traditional" MARC records to be automatically transformed into "FRBRised" records. Three major OCLC realisations are based on this FRBRisation algorithm: xISBN, Open WorldCat, and FictionFinder.

xISBN is a service that supplies the ISBNs that identify the various editions of a single Work. When you submit it an ISBN, by entering the URL http://labs.oclc.org/xISBN/ immediately followed by an ISBN, it returns a list of associated ISBNs. Perhaps you will wonder what you can do with a raw list of ISBNs? As a matter of fact, OCLC's xISBN service is intended to be used by machines, rather than people. If you, as human beings, cannot do much with just a raw list of ISBNs, machines can do very interesting things with it.

For instance, if you launch on Google or Yahoo a query on the words "maalouf crusades," together with the phrase (within quotation marks) "find in a library" (or, alternatively, the instruction site:www.worldcatlibraries.org), the first hit you get is a WorldCat bibliographic record describing one of the many editions of Amin Maalouf's *The crusades through Arab eyes.* Under that bibliographic record, you will notice the word "Editions," which carries an active link. When you click on it, you get the list of the bibliographic records for all the various editions of

Amin Maalouf's Work, in various linguistic Expressions, that are available in WorldCat. You can pick one reference in that list, and thus obtain the display of a record for a distinct edition.

OCLC's prototype called FictionFinder (http://fictionfinder.oclc.org) shows how the display of bibliographic lists for textual works of fiction can be simplified by using the FRBRisation algorithm. For instance, the various editions of Najib Mahfuz' novel *Midaq alley* are arranged by language, under a brief record at the Work level. By clicking on an individual entry in that list of languages, you get information about the editions of the Work in a given language, and complete records for individual editions. [OCLC's Research Office also has a new experimental prototype called Curiouser for even greater use of the FRBR concepts to bring together the manifestations and expressions of works in very user-friendly displays.]

RLG

RLG's OPAC "RedLightGreen," which can be accessed at http://www.redlightgreen.com, reduces the four FRBR levels to only two: Work, and Manifestation. Unfortunately, due to inconsistencies in cataloguing and various degrees of accuracy in older bibliographic records, it is not always possible to cluster automatically all Manifestations under the Work they are of (by the way, this is true for WorldCat and FictionFinder as well). For instance, if you submit the words "diwan abu nuwas" to RedLightGreen, what you get is not just one uniform title under which all editions of Abu Nuwas' poetry would be clustered. But you get subsets of editions, for example by clicking on the first line in the hit list you get "9 editions published between 1958 and 2002 in 2 languages." By clicking on the title of the Work, you get information about one of its editions. And by clicking on the phrase "9 editions," you get the complete list of those 9 editions, from which you can pick individual records. One advantage to RedLightGreen is its ability to offer the user suggestions of controlled forms of names and subjects to search and to transform selected citations into whatever citation style the user may want from a list of standards, like the Chicago Manual of Style.

Current evolution of FRBR

FRAR (Functional Requirements for Authority Records)

FRBR focuses on the content of *bibliographic* records only (as opposed to authority records), and access points to bibliographic records, but does not go into detail about authority records. This is the reason why the FRANAR (Functional Requirements and Numbering for Authority Records) Working Group was created in 1999 under the joint auspices of the IFLA Division of Bibliographic Control and the Universal Bibliographic Control and International MARC Programme (the late UBCIM). The first of the three terms of reference for the FRANAR Group was "to define functional requirements of authority records, continuing the work that FRBR initiated." Tom Delsey proved as instrumental in designing the FRAR model as he had been in designing the FRBR model. "Person" and "Corporate Body", that were only represented in FRBR by a heading, are now fully modelled, and the entity "Family" has been added in recognition of its importance to the archival community.

FRSAR (Functional Requirements for Subject Authority Records)

The peculiar nature of subject relationships is only alluded to in FRBR, and is not dealt with in depth in FRAR. The IFLA Division of Bibliographic Control formed therefore in August

2005 a third Study Group, which is in charge of defining an extension of the FRBR/FRAR model in order to account for classification and indexing. This new model will be called FRSAR (Functional Requirements for Subject Authority Records). With FRBR, FRAR, and FRSAR, we will have at last a complete model for the information we put in library catalogues.

FRBR Review Group

FRBR is not meant to remain such as it stands forever and ever. It was decided in 2003 to go through a reviewing process, and to form a FRBR Review Group affiliated to IFLA. Additionally, another working group was formed in 2005, in order to assist the FRBR Review Group in exploring the conceptual difficulties in modelling "aggregates," that is: Manifestations that embody more than one Work, continuing resources, multi-volume monographs, and other whole/part issues…

FRBR and the Meeting's 5 "Focus Topics" (that happen to be 6…)

Tomorrow, you will be invited to work on "Focus Topics" that have been defined prior to the Meeting. At the first two IME ICC meetings there were 5 working groups: Personal names, Corporate bodies, Uniform titles and GMDs, Seriality, and Multipart structures (composites versus aggregates). There are actually 6 of them, since the GMD issue and the uniform title issue are gathered under one Focus Topic. These 6 topics can be organised into three broad categories: "Appellations," "Types," and "Aggregates." Let us have a quick glance at each of them in light of FRBR.

"Appellations"
Names of personal authors
The attributes defined by FRBR for the Person entity distinguish between the "name" of a person and that person's "dates," "title," and "other designation;" actually, these are all the elements that make up the heading for a person in a bibliographic record, and it surely would have been enough, for the purposes of FRBR, to define just one attribute: "heading." The FRAR model addresses some issues that were out of the scope of FRBR: Is an instance of the Person entity supposed to be an actual person in the real world, or "something" else, and what? I mean: Can a real person be represented by two instances of the Person entity; inversely, can two real persons be represented in the catalogue universe by only one instance of the Person entity? The FRAR model, relying on AACR2, defines the notion of *bibliographic identity,* that is: that intermediate between the real world and the catalogue universe. This is actually what the Person entity is about. A real-life person may have several bibliographic identities (as in the case of pseudonyms), and several distinct real-life persons may be merged into one single bibliographic identity (as in the case of families and shared pseudonyms, but also in the case of undifferentiated names). An instance of the FRBR Person entity does not necessarily "overlap" with what we usually understand as "a person" in the real world.

In most cases we strive to "control" bibliographic identities, that is: we strive to be quite aware of when distinct bibliographic identities correspond to one real-world person (in which case we would like to see links between bibliographic identities), and when one bibliographic identity corresponds to distinct persons. Families and shared pseudonyms *can* be controlled; undifferentiated names of persons are grouped together and not separately distinguished. The question therefore is: To what extent is that lack of control tolerable? Can it be tolerated at all?

Does it have a substantial impact on the practicability of our catalogues for our users? Do they complain about it? Is the notion of "bibliographic identity" (as distinct from the notion of "real-world person") helpful or confusing for end-users? It certainly strives to use the name the person used on manifestations, which is what an end-user would see at a bookstore or on a borrowed book from a friend or a library – putting the user first in terms of what they are likely to know. There would, of course, be references from variant forms of names.

Names of corporate bodies

This is very much the same issue. The FRBR attributes for the Corporate Body entity actually could have been replaced, for the specific purposes of FRBR, with only one attribute, "heading." Here again, the Corporate Body entity does not correspond to a real-world corporate body, but rather to the notion of "bibliographic identity" as defined by FRAR and AACR2. Does every name change reflect a transformation of a corporate body into a new corporate body? Should every name change result in the definition of a new bibliographic identity, or should all name changes be recorded as cross references for the same bibliographic identity, as end-users may be more responsive to the notion of a corporate body's continuity over time, than to its occasional modifications? Should we go on regarding congresses as "corporate bodies," while they are closer to the FRBR notion of "Event?" Many of these old concepts are evolving as we explore these models.

Uniform titles

Title attributes are defined in FRBR at three levels: Work, Expression, and Manifestation. Appendix A in the *FRBR Final Report* explicitly states that the title of a Work may be either a uniform title or the title proper; that there currently is no prescription at all for the title of an Expression; and that the title of a Manifestation may be the title proper, a parallel title, a variant title, a transliterated title (all of them are *transcribed* titles), or a key title (which actually is a title *created* by cataloguers).

The title proper can therefore relate to both the Work and the Manifestation entities, while we do not know from FRBR what a title of an Expression could be.

I think it is arguable and sensible to state that the title of an Expression actually consists of the title of the Work realised by the Expression, plus any combination of additional elements taken among all of the other attributes defined for the Expression entity. The absence of any prescription as to what constitutes the title of an Expression indicates that we currently have no identification device at the Expression level. Although many uniform titles already provide for elements that identify an expression, there is no system, in library catalogues, for consistently citing or referring to specific translations or versions of a given textual work, specific recorded performances of a given musical work, specific states of a given engraving, etc. Should our future International Cataloguing Code contain provisions for such a system? If so, which elements, and in which order, are strictly indispensable in order to cite/refer to a specific Expression, either in the role of title heading or in the role of subject heading?

"Types"
GMDs (General Material Designators)

GMDs are not mentioned in the *FRBR Final Report.* One reason is that what we now know as "GMDs" through the ISBDs are a mixture of terms expressing content and others expressing the carrier for that content – information for the expression level (content) and for the manifestation level (carrier). This is a fact of historical convenience, but is very illogical. Perhaps

FRBR lacks a "type" attribute for each of the three upper entities: Work, Expression, Manifestation. Perhaps what we have in mind when we talk about GMDs would be a combination of these 3 "type" levels, such as, for instance (these are only suggestions, I am aware they are not quite consistent):

textual work – expressed as sound – on physical carrier

textual work – expressed as written word – on manuscript

textual work – expressed as written word – on microform

musical work – expressed as notation – on printed material

musical work – expressed as sound – in an electronic resource on line

In some cases the expression level might be omitted:

cartographic work – in an electronic resource on line

motion picture – on physical carrier

multimedia work – on physical carrier(s).

The problem is that we would like GMDs to be as concise as possible. All these suggested GMDs are much too long.

"Aggregates"
Seriality issues of when to make a new record

Once again, the issue in cataloguing is about the correlation between appellations and the intrinsic essence of the entities that bear those appellations. Does every title change make a distinct, new Work out of a continuing resource? Does a title change affect the Manifestation level only, or the Work level of a continuing resource? Does the key title identify a continuing resource as a Work, as an Expression, or as a Manifestation? Why is the name of the originator of a continuing resource integrated into a key title as a qualifier, whereas for other kinds of works we most often create author-title headings? Why don't we deal with key titles in authority records rather than bibliographic records? Why do we make bibliographic records for continuing resources, and authority records for trademarks, that are very much akin to continuing resources? Is it justified at all to create *bibliographic* records for periodicals? There have already been proposals to explore using authority records instead, so the FRBR model has also generated more thinking about these concepts.

Multipart structures

It often happens that *one* publication contains *more than one* individual Work. Is the purpose of library catalogues to enable end-users to identify and retrieve a given publication as a whole, or to identify and retrieve a given Work in all of its possible Manifestations? What is the basic unit of bibliographic information: the Work, or the Manifestation? This issue was known in the 1960s as the debate between "literary unit" (i.e., the Work) and the "bibliographic unit" (i.e., the Manifestation).

In a recent lecture at the University of Illinois,[1] Barbara Tillett said that "today, this debate is no longer relevant," as "we don't need to prefer bibliographic unit or literary unit over the other, because our bibliographic and authority information can be rearranged by computer systems to display the elements that we have included in description and access points." I totally agree with her that the debate is no longer relevant and the current (and future) possibilities of our

---

[1] TILLETT, Barbara B. *Cataloging for the future* [on line]: delivered as the 2004 Phineas L. Windsor lecture at the University of Illinois Graduate School of Library and Information Science, October 13, 2004. [Urbana-Champaign, Ill.]: [GSLIS Publications Office], [2004] [cited 12 September 2005]. Available from the Internet: <http://puboff.lis.uiuc.edu/catalog/windsor/windsor_tillett.html>.

computer systems contribute to solve the problem, and yet the issue still has to be mentioned in this series of IME ICC meetings, because the way national cataloguing codes and individual library practices deal with this issue is so inconsistent that the same publication is often assigned very different descriptions in different catalogues.

For instance, let us consider the following publication: *Récits en noir et blanc : cinq nouvelles arabes* [Qiṣ aṣ bi-al-abyaḍ wa-al-aswad] ("Narratives in black and white: five Arab short stories"), issued in Paris in 2001. The five Arab short stories contained in this publication were selected and translated into French by François Zabbal, who also wrote a foreword for the entire collection. Should we regard the abstract, textual content of "Narratives in black and white" on the whole as a Work, then? And is François Zabbal the "creator" of that Work? However, this is not the issue I intend to address here, as collections and anthologies are one of the many tricky topics that the IFLA Working Group on Aggregates will have to solve. My main concern here is: How to deal with the individual, distinct *Works* embodied in such a publication? In addition to Zabbal's foreword, which is a Work in itself, "Narratives in black and white" contains the following Works by major Arab writers, both in their original Arabic Expression and in Zabbal's French Expression: *Le rêve* [al-Ru'yā] ("The dream"), by Abdessalam al-Ujayli; *Le chanteur de nuit* [Mughannī al-layl] ("The night-singer"), by Zakaria Tamer; *Le duel* [al-Mubārazah] ("The dual"), by Muhammad el Bisatie; *Le funeste oiseau d'or* [Tā'ir al-nahs al-dhahabī] ("The ominous golden bird"), by Ibrahim al-Koni; and *Ochba* [Ushbah], by Salma Matar Saif. Depending on which cataloguing code and which cataloguing policy we apply, we can:

– just ignore all of them, as there are more than three;
– mention them, possibly in square brackets, as an "other title information" statement;
– mention them in a contents note, without indexing them;
– mention them in a contents note, *and* create author/title added entries for all of them to be retrievable;
– create analytical records.

What should we recommend in an International Cataloguing Code?

Conclusion

It would be too wonderful if FRBR solved all theoretical problems, or could help us solve all theoretical problems. This is obviously not the case. But FRBR provides us with a conceptual tool that is very helpful in assessing and questioning the way we catalogue. The clear distinctions it makes between what is "physical" and what is "abstract" in the "things" we describe, and the various levels it identifies in the "content" of "publications," are useful to keep in mind when we think about how to improve our catalogues and how to improve the services we provide our end-users with. Those distinctions should definitely be reflected in our future International Cataloguing Code.